Psychosocial Risk Factors for Upper Respiratory Infections: ASSESSMENT OF UPPER RESPIRATORY ILLNESS DURING BASIC TRAINING*

Ross R. Vickers, Jr.

Linda K. Hervig

Applied Physiology Department
Naval Health Research Center
P.O. Box 85122
San Diego, CA 92138-9174

*Report No. 88-23 was supported by the Naval Medical Research and Development Command, Navy Medical Command, Department of the Navy under Research Work Unit MR04101.07-6004. The views presented are those of the authors and do not reflect the official policy of the Department of the Navy, Department of Defense, nor the U. S. Government.

SUMMARY

Upper respiratory illness (URI) is common in military populations. The prevalence, cost, and antecedents of URI cannot be accurately estimated without considering infections which can affect performance but do not lead to seeking medical care. Symptom reports have served this purpose in civilian settings, but available measures may be inappropriate for individuals performing physically demanding work in psychologically stressful situations. This study developed a symptom report URI scale suitable for studies in such circumstances.

A symptom checklist was constructed with thirty-nine possible URI symptoms identified from prior experimental, epidemiological, and clinical studies of viral infections and URI syndromes. The checklist also included two symptoms of musculoskeletal problems and four symptoms of other minor non-URI illnesses. Four items assessed subjective attributions regarding the causes of symptoms, including allergic reactions, common cold, flu, and muscle strains and sprains. This checklist was administered to four samples of Navy recruits at approximately weekly intervals during basic training. This setting was selected because basic training combines exceptional psychological and physical demands with substantial exposure to viral pathogens.

Symptoms were evaluated as URI indicators by comparing recruits with a common cold, but no other illness, to recruits reporting no illness, and to recruits reporting musculoskeletal problems, but no other illness. Twenty-five symptoms did not reliably discriminate between the groups. Six symptoms which did reliably discriminate were more pronounced in recruits with musculoskeletal problems than in recruits with URI. The remaining 14 symptoms were more pronounced among recruits with URIs, but two were eliminated because they generally were only slightly higher in the URI group. Four other symptoms were deleted to minimize item redundancy. These exclusions left sore throat, productive cough, stuffed-up nose, dry cough, sinus pain, sneezing, fever, and hoarseness as symptomatic URI indicators.

Analyses showed that the eight symptoms retained defined a continuum of illness severity. Adjustments for allergic reactions and musculoskeletal problems were shown to be necessary to ensure that URI measurement was not contaminated by other illnesses. The resulting URI measure combined acceptable measurement precision with the moderate temporal stability

expected of a transient illness. The URI measure correlated with general symptom reporting, but the association was low enough to ensure that the URI composite was not merely an index of somatic stress reactions or hypochondriasis.

The eight-item composite provides a URI measure suitable for epidemiological surveillance of military populations. The most notable difference between this composite and similar measures used in civilian studies is the absence of general malaise and muscle aches and pains. In military populations, these symptoms would confuse URI with musculoskeletal problems. The proposed URI measure has face validity, is consistent with the clinical practice of ruling out other illnesses when diagnosing URI, and has good measurement precision. However, interpretation of results obtained with this measure should allow for potential biases arising from general symptom reporting tendencies.

Table of Contents

List of Tables

INTRODUCTION

Upper respiratory infections are a common health problem in the military. Typically, the occurrence of such infections is assessed by data concerning hospitalization or outpatient treatment. However, many people do not seek medical care, even when respiratory infections significantly restrict their activities (e.g. National Center for Health Statistics, 1983, 1986). Estimates derived from medical care records, therefore, can significantly underestimate the substantial costs associated with URI (Harlan, Murt, Thomas, Lepkowski, Guire, Berki & Landis, 1986). A simple method of achieving more complete surveillance of upper respiratory infections in military populations is desirable to obtain more meaningful estimates of the costs of URI in these populations. This paper describes the development of a respiratory symptom report measure suitable for surveillance of upper respiratory infections in general military populations.

Symptom reports have been used to assess upper respiratory infections in experimental and epidemiological studies (Gwaltney, Hendley, Simon & Jordan, 1967; Jackson, Dowling, Spiesman & Boand, 1958; Beare & Reed, 1977; Monto, Napier & Metzner, 1971). Symptom reports are easily obtained, inexpensive, and correlate with viral shedding (Forsyth, Bloom, Johnson & Chanock, 1963; Totman, Kiff, Reed & Craig, 1980) and clinical URI evaluations (Roden, 1958, Totman, Reed & Craig, 1977). Although these previously reported correlations illustrate the validity of symptom reports, such reports undoubtedly are imperfect indicators of pathological processes. For this reason, the acronym "URI" will refer to "upper respiratory illness" in the remainder of this paper to distinguish illness, the subjective assessment of well-being, from disease, the presence of pathological conditions that are one determinant of illness.

Available URI measures may not be suitable for military personnel who frequently perform heavy physical work under extreme environmental conditions while separated from their families and living and working in crowded conditions. These factors, which are not typical of the general United States population, may produce trauma, musculoskeletal injury, and psychological distress. As a result, some symptoms commonly included in URI measures (e.g., malaise, muscle aches, headaches) may confound URI with other physical and psychological problems when employed in military

populations.

Despite potential problems, appropriately constructed symptom assessments can measure URIs effectively in military populations. Prior research in military basic training has shown that symptom reports correlate with viral shedding (Forsyth, et al., 1963), antibody responses (Arlander, Pierce, Edwards, Peckinpaugh & Miller, 1965; Lytle & McNamara, 1967) and serum immunoglobulins (Lytle, Rytel & Edwards, 1966; Vickers, Hervig, Edwards & Rahe, in preparation). Unfortunately, the number of symptoms evaluated and the weights assigned to specific symptoms has varied in these studies. Also, no attempt has been made to evaluate the impact of other illnesses or general symptom reporting tendencies on these symptom measures. The result is that there is no accepted standard assessment for URI in military personnel. With these points in mind, the objectives for the present study were:

(a) Identify symptoms suitable for measuring URI in a military population facing significant physical and psychological stresses.

(b) Describe the psychometric characteristics of the resulting URI measure.

(c) Evaluate the influence of other illnesses and general symptom reporting on the URI measure.

METHOD

Sample

Health data were collected from four samples of Navy recruits who gave informed consent when asked to participate in studies to identify predictors of URI (Table 1). Recruit training was the setting for study, because URIs occur frequently and because recruits face significant physical and psychological stresses. Analysis of variance indicated significant sample differences in age ($p < .001$), despite the small absolute magnitude of the differences. Chi-square analyses indicated significant sample differences in racial composition ($p < .002$) and educational level ($p < .001$). Log-linear analysis (Brown, 1983) indicated that the significant chi-squares were attributable largely to the low proportion of Whites and the high proportion of recruits without diplomas in Sample A. Given the heterogeneity of the samples, tests of the homogeneity of findings across samples were included in all analyses to determine whether the sample differences influenced study results.

Table 1

Demographic Characteristics of the Samples

		Sample A	Sample B	Sample C	Sample D
Age					
	Mean	18.82	19.06	19.51	19.31
	S.D.	2.24	2.49	2.86	2.63
	Range	16-35	17-33	16-33	17-33
	N	681	596	552	591
Race					
	Hispanic	8%	4%	7%	7%
	Black	21%	14%	16%	17%
	White	65%	78%	71%	70%
	Other	7%	5%	6%	5%
Education					
	No Diploma	13%	4%	5%	3%
	G.E.D.[a]	5%	3%	3%	2%
	H.S. Diploma	83%	93%	92%	95%

[a]G.E.D. = Graduate Equivalence Diploma

Health Status Measures

Forty-five symptoms were included in the assessment of health status. Of these, 39 were symptoms which had been identified as potential URI indicators from prior reports of experimentally-induced viral infections, epidemiological studies of URI, and clinical descriptions of URI syndromes (Anderson, Patriarca, Hierholzer & Noble, 1983; Dingle, Badger & Jordan, 1964; Dowling & Lefkowitz, 1963; Forsyth, et al., 1963; Gwaltney, et al., 1967; Jackson, et al., 1958; Jackson & Muldoon, 1973). Items concerning symptoms of skin irritation, sunburn, blisters, and constipation were added to assess the general tendency to report minor health problems. A similar composite was a weak URI correlate, but a strong correlate of hypochondriacal tendencies in a previous study (Vickers, Hervig & Edwards, 1986). If this prior finding could be replicated, it would rule out general symptom reporting tendencies as a factor affecting URI reports. The remaining two symptoms concerned muscle aches and muscle cramps, common musculoskeletal problems in basic training.

The health status assessment also included four items corresponding to Verbrugge's (1986) concept of an attributed cause of symptoms. These items asked about perceptions of four general illness syndromes, the common cold, the flu, musculoskeletal injury, and allergy. The common cold was included, because this attribution can help distinguish acute URI from chronic respiratory symptomatology (Jackson, et al., 1958). The flu attribution was included to determine whether many URI symptoms occurred among people who described their general illness as flu rather than a cold. The allergy attribution made it possible to follow the clinical practice of ruling out allergy as a source of URI symptoms before diagnosing an illness as a cold (Lowenstein & Parrino, 1987). The musculoskeletal item represented the most frequent non-URI illness in basic training. Evaluating the effect of this health problem on URI provided a check on a potentially significant source of contamination if other illnesses influence URI reports.

Data Collection

Symptom checklists were completed at seven data collection sessions. These sessions were conducted 4, 12, 19, 26, 37, 46, and 53 days after beginning training for approximately 50% of the participants. The sessions were conducted two days later for the remaining participants because a weekend intervened between the start of the study and the fourth day of the

Table 2

Selection of URI Symptoms

Symptom	Analysis of Variance: Coefficient of Concordance	Number Sig	URI versus Musculoskeletal: Average t	Combined z
Final URI Item				
Dry Cough	1.000	14	2.38	9.28
Stuffed-up Nose	.941	16	4.30	17.38
Sore Throat	.891	16	2.83	11.23
Sneeze	.824	15	2.02	8.40
Productive Cough	.785	16	3.80	14.89
Sinus Pain	.774	11	1.51	5.65
Hoarseness	.684	13	2.00	8.43
Fever	.660	15	1.77	7.11
Redundancy Criterion				
Runny Nose	.813	16	4.87	19.47
Sputum	.848	15	2.56	10.05
Scratchy Throat	.848	14	2.70	10.67
Burning Throat	.543	9	1.40	5.61
Small Average *t-test* Value				
Chills	.766	9	0.70	2.72
Headache	.707	12	0.80	3.19
Musculoskeletal Criterion				
Muscle Aches	.848	16	-4.54	-17.81
Aching Joints/Bones	.848	14	-2.90	-11.39
Muscle Cramps	.813	16	-4.31	-16.88
Stiff Neck	.777	10	-1.63	-6.37
Sweating	.738	13	-0.95	-3.72
Malaise	.750	14	-0.18	-0.26
Weak ANOVA Differences				
Trouble Swallowing	.757	7		
Chest Pain	.687	5		
Irritability	.672	6		
Trouble Sleeping	.672	6		
Watery Eyes	.663	8		
Abdominal Pain	.622	2		
Blisters	.613	3		
Shivering	.609	5		
Loss of Appetite	.578	7		

NOTE: Lump in the throat, wheezing breath, constipation, trouble hearing, nausea, dizziness, sunburn, cold sweats, skin irritation, earache, vomiting, painful breathing, cold sores, irritated eyes, rash, and diarrhea failed the first criterion.

training schedule for these participants. This schedule was the closest possible approximation to a weekly assessment within the constraints of the training schedule.

At each session, recruits indicated the severity of each symptom over the preceding three days of basic training. Severity was indicated by marking the appropriate space on an optical scanning sheet with response options ranging from "Not at All Severe" (1) to "Extremely Severe" (5). Items were read aloud to eliminate any effects of reading problems and to encourage recruits to ask for definitions and elaboration if needed. The instructions successfully stimulated questions, as requests for repetition, explanation, and definitions were common, particularly during the early administrations of the checklist.

RESULTS

Symptom Selection

Subjective Health Status and Symptoms. An ideal URI symptom would be present only in individuals with URI. To determine how well symptoms approximated this ideal, three groups of recruits were identified on the basis of their subjective attributions regarding cold, allergy, and muscle strain: (a) No Illness; (b) Musculoskeletal Problems Only; and (c) Cold Only. Recruits who claimed more than one health problem (e.g., cold and muscle strain) or who reported an allergic reaction, but no other illness, were excluded. These exclusions were judged reasonable, because the source of reported symptoms would be ambiguous whenever multiple health problems were present.

Sixteen (four days for four samples) one-way analyses of variance (ANOVAs) were performed for each symptom. These analyses compared the three groups for sessions 1 through 4 for each sample. Group classification for each ANOVA was based on illness attributions for that session.

The general criterion for evaluating symptoms was that the URI group should consistently score significantly higher than the other two groups. Any symptom satisfying this general criterion must produce a consistent pattern of group differences. How well each symptom met this minimum requirement was evaluated by ranking the mean scores for the three groups

for each of the 16 analyses. Kendall's coefficient of concordance, W (Hays, 1963, pp. 656-658), then was computed for the set of 16 rankings. Sixteen symptoms which produced coefficients less than .500 were excluded from further consideration.

The ordering of group means could be consistent even if the differences between groups were too small to be of value in distinguishing between groups. The F-tests from the ANOVAs provided a basis for evaluating the size of the group differences, because this statistic compares observed variation in group means to the estimated variation that would be expected by chance. The associated significance test determines whether the group differences are large enough that they cannot reasonably be attributed to chance. Nine of the 29 symptoms which produced stable rank orders of group means produce statistically significant ($p < .05$) differences in fewer than 9 of the 16 analyses (Table 2) and were dropped from further consideration.

It was anticipated that several symptoms would be more prevalent in the musculoskeletal group than the URI group, so the symptom scores for the musculoskeletal group were compared to those of the URI group. These comparisons produced 16 t-tests for each of the 20 symptoms remaining at this point in the analysis. The method of adding t-tests (Rosenthal, 1978) was used to combine the findings from the 16 analyses into an overall evaluation (Table 2). Six symptoms produced average t-scores less than zero, indicating higher overall scores in the musculoskeletal group. Two additional symptoms, "chills" and "headache", were eliminated from further consideration even though they were more pronounced in the URI group. This deletion was based on the observation that the typical differentiation between groups for these symptoms ($z < 3.20$) was substantially less than that of the remaining 12 symptoms ($z > 5.61$).

Screen for Redundant Symptoms. Symptoms forming a clinical syndrome should be correlated but not so highly correlated that it could be inferred that any two items measure a single symptom. Including multiple items for a symptom would give this symptom undue weight in measuring URIs. Therefore, correlations between the remaining 12 symptoms were examined for the data for sessions 1 through 4. One item was eliminated from any pair that produced more than one correlation exceeding $r = .60$, provided the correlations exceeding this cutoff were not obtained in a single sample. Symptom pairs exceeding the criterion were sputum and productive cough

(r > .60 14 times, median r = .70), sore throat and burning throat (r > .60 8 times, median r = .60), sore throat and scratchy throat (r > .60 twice, median r = .53), burning throat and scratchy throat (r > .60 7 times, median r = .59), and runny nose and stuffed-up nose (r > .60 10 times, median r = .62).

Productive cough, sore throat, and stuffed-up nose were retained for the URI measure. Productive cough was selected in preference to sputum, because it is a less technical term. Sore throat was selected in preference to scratchy throat and burning throat, because it produced better discrimination between the URI and musculoskeletal groups (Table 2) and has been more frequently used as a URI indicator than either of the alternative symptoms in prior research. Stuffed-up nose was selected in preference to runny nose, because recruits undergo a gas chamber exercise during training. This exercise produces extensive mucous flow, so including runny nose would artificially increase URI scores at that point in training. The deletions clearly eliminated items with exceptionally high correlations as only 3 of the 436 correlations between the eight remaining symptoms were as large as .60.

Psychometric Evaluation

Dichotomy or Continuum? Common experience suggests that URIs vary with regard to number and severity of symptoms. Despite this common experience, symptom reports often are used simply to classify people into categories of healthy and ill, rather than to determine differences in location on a continuum of illness severity. Meehl's (1973) MAXCOV test for discrete classes, therefore, was applied to the symptom reports as a preliminary to forming a symptom composite to assess URI. The results, summarized in Appendix A, indicated that the URI reports should be scored to represent locations on a continuum of disease severity rather than to assign respondents to a dichotomous classification. All subsequent analyses assumed the existence of a continuum of illness severity.

Dimensionality of the URI Symptoms. Even if symptom reports define locations on a continuous severity of illness dimension, symptoms could occur in unique combinations that defined distinct illness syndromes. If so, URI symptoms could be classified into subsets with higher correlations within subsets than between subsets. Principle factors analyses (cf., Gorsuch, 1984) were conducted to test for this possibility, because this

procedure would produce evidence for multiple factors if the pattern of correlations expected with multiple syndromes were present in the data.

The first factor was large in each of the 16 factor analyses (4 days for four samples) with a median proportion of variance explained of 50.3%. No analysis produced mathematical evidence for more than one factor (i.e., more than one eigenvalue > 1.00). Each symptom was strongly related to the first factor, as indicated by average factor loadings well above the .30 criterion frequently used to define significant markers for factors (Table 3). The small standard deviations for the factor loadings across analyses showed that the average loadings were not obtained by combining some extremely high loadings with other relatively low loadings. The small variance also implied that sample differences were minor. Thus, the items defined a single syndrome, and all of the symptoms were pertinent to that syndrome.

Table 3

Factor Loadings for Individual Symptoms

Symptom	Mean[a]	S.D.[a]
Dry Cough	.64	.07
Sore Throat	.77	.03
Sneeze	.63	.03
Stuffed-up Nose	.77	.02
Productive Cough	.73	.04
Sinus Pain	.65	.03
Hoarseness	.71	.03
Fever	.69	.04

[a] The mean and standard deviations for the factor loadings were computed from loadings obtained in sixteen factor analyses. one for each of four study days for each sample.

Psychometric Assessment of URI Composites. An ideal URI measure would assess individual differences in illness severity with high precision at a given point in time but would have moderate stability across time. Moderate cross-time stability is predicted from evidence that URIs typically last three weeks or less (Roden, 1963). An ideal URI measure also would be unaffected by individual differences in willingness to report symptoms, because greater effects of general symptom reporting imply less precise assessment of the specific illness of interest. Low specificity implies increased risk that empirical associations between the URI composite and other variables would reflect associations to general symptom reporting, rather than a true association to URI. Appendix B describes the measure of general symptom reporting used in these analyses.

Heise's (1969) procedure for estimating measurement precision and temporal stability was applied to estimate these psychometric aspects of the URI and General Symptom measures. This procedure provided estimates of both aspects of measurement and could be applied to general symptom reporting. Alternative procedures based on internal consistency criteria (e.g., Cronbach's alpha) were logically inappropriate for the latter variable. The reliability estimates from Heise's (1969) procedure were used to correct URI-General Symptom correlations for attenuation by measurement error (Nunnally, 1978, pp. 219-220). Without this correction, scales with poorer reliability might be regarded as less contaminated, merely because lower reliability reduced their observed correlation to general symptom reporting.

Two scoring methods were employed to create symptom composites for these analyses. The first composite (hereafter URI Severity) averaged the reported severity for the 8 URI symptoms. The second composite (hereafter, URI Symptoms) recoded the raw data to indicate merely whether the symptom was present or absent, then counted the number of symptoms reported. This comparison was made to determine the effects, if any, of stylistic differences between individuals with regard to the use of the severity rating scale on the resulting URI reports.

Data from the initial session were excluded from these analyses, because scores for that session were atypical with regard to the temporal stability of the scales (Appendix B). Including data from these sessions would have biased the analysis results, because this deviance would affect all of the coefficients to be estimated.

Table 4

Comparison of Alternative URI Measures

	Symptom Severity	Symptom Count	Adjusted Severity
Early Period			
Reliability	.843	.759	.798
Stability			
1-2	.671	.709	.617
2-3	.728	.810	.662
Correlation to General Symptoms	.582	.566	.516
Late Period			
Reliability	.744	.716	.669
Stability			
5-6	.833	.828	.755
6-7	.826	.873	.779
Correlation to General Symptoms	.746	.755	.610

NOTE: Table entries are averaged correlations for Samples A-D computed using Fisher's r-to-z transformation (Hays, 1963, p. 529-533). See Appendix C for analysis details.

The first two columns of Table 4 provide the coefficients necessary for the psychometric comparison of the URI Severity and URI Symptom approaches. The coefficients reported generally showed statistically significant variation between samples, but the relative magnitudes of the coefficients were consistent across the four samples studied (Appendix C). The figures in Table 4, therefore, were judged satisfactory for the purpose of comparing alternative measures, because the relative size of the coefficients was the important consideration. Recalling that an ideal URI scale would combine high measurement precision, labelled reliability in the table, with low to moderate temporal stability, and little or no association to general symptom reporting tendencies, the important comparisons between URI Severity and URI Symptoms can be summarized as follows:

Reliability: Severity > Symptoms
Stability: Symptoms > Severity
U-G Correlation: Severity = Symptoms

URI Severity was preferable to URI Symptoms with regard to two of the three criteria and roughly equivalent for the third. URI Severity also can be preferred to URI Symptoms on the basis that this measure employs participants' actual responses rather than a transformation of those responses. For these reasons, URI Severity was retained as the measure of choice for subsequent analyses.

The difference in the estimated coefficients for data from early and late sessions was notable. This difference may be related to the fact that the rate of infections is much higher early in basic training (Edwards & Rosenbaum, 1971). Actual disease, therefore, should be a more important contributor to overall variation in URI scores during that period, so it is plausible that URI measures would approximate an ideal pattern more closely early in training.

Effects of Non-URI Illnesses. Many recruits reported allergies which could bias URI estimates by producing respiratory symptoms and/or musculoskeletal problems which conceivably could alter general symptom sensitivity and reporting. The impact of these concurrent illnesses on URI Severity was assessed by classifying recruits as having or not having (a) a cold, (b) an allergy, and (c) a muscle strain on the basis of their attributions for illness. Dichotomous classifications were used to construct cross-products representing potential statistical interactions between illnesses (Cohen, 1978). Stepwise multiple regressions entered "cold" first, then "allergy", then "muscle strain", and, finally, the four interaction terms as a group. Applying this approach to the data from sessions 2 through 4 for each study showed:

(a) The four interaction terms combined explained an average of 0.8% (range = 0.2% - 2.4%) of the URI variance.
(b) The "Allergy" main effect explained an average of 5.8% (range = 3.3% - 9.7%) of the URI variance. The increment in variance explained was highly significant ($p < .001$) in all 12 analyses.
(c) The "Muscle Strain" main effect explained an average of 4.5% (range = 1.8% - 6.7%) of the URI variance. The increment in

variance explained was highly significant ($p < .001$) in all 12 analyses.

(d) Analyses of covariance with the "allergy" dichotomy as the group classification variable and "cold" as a covariate showed that, on the average, the reported URI score was 0.49 points higher if the recruit reported an allergy than if he did not. A similar analysis with "cold" and "allergy" as covariates showed that the average URI score was 0.29 points higher if a recruit reported a "muscle strain" than if he did not.

Adjusted URI Severity scores were created by reducing the URI severity score by 0.49 points if an allergy was reported and 0.29 points if a muscle strain was reported. These scores correlated highly with the raw score (median r = .943; range = .935 - .946). Psychometric evaluation of the Adjusted URI Severity score showed that this composite had slightly lower measurement precision than URI Severity, but also produced slightly lower temporal stability, and substantially lower overlap with general symptoms (r = .52 versus .58 for sessions 2-4, .61 versus .75 for sessions 5-7; Table 4).

Adjusted URI severity was judged preferable to URI Severity. The adjustments help rule out competing interpretations of the URI composite. The lower correlation to General Symptom reporting reduces the likelihood that significant associations between Adjusted URI Severity and other variables will prove to be indicators of the relationship between those variables and any general tendency to report somatic symptoms in response to stress. Adjusted URI Severity also is consistent with the clinical practice of ruling out other possible types of illness when making a diagnosis of upper respiratory infection (Lowenstein & Parrino, 1987). These gains from incorporating adjustments for allergy and musculoskeletal problems seemed adequate to offset the minor loss in measurement precision represented by the lower reliability of the Adjusted URI Severity score.

Assessment of Cumulative URI Experience. Average scores for the adjusted URI composite were computed for the early and late periods of training and for the entire course of training. These averages assessed the cumulative severity of illness and provided a reasonable indicator of individual differences in illness during basic training.

Correlations between the cumulative illness composites were comparable

across the four samples, allowing for the fact that 16 significance tests will produce at least one statistically significant ($p < .05$) result by chance 54% of the time. Therefore, correlations were averaged across the four samples to provide an overall description of the associations (Table 5).

URI scores were much less stable over time than General Symptom scores (.46 versus .66). This comparison provides a frame of reference for asserting that the URI measure has moderate temporal stability, a desirable attribute for a measure of a transient illness. The URI-General Symptom correlations were moderately large but low enough to indicate that the URI composite did not merely measure hypochondriacal tendencies or acute psychological reactions to basic training. Nevertheless, these correlations were large enough to make it important to take general symptom reporting into account when interpreting results obtained with URI symptom reports.

Table 5

Correlations Between Aggregate Upper Respiratory Illness and General Symptom Report Measures

	URI			General Symptoms		
URI	Total	Early	Late	Total	Early	Late
Total	1.00					
Early	.90[a]	1.00				
Late	.80[a]	.46	1.00			
General Symptoms						
Total	.54	.48	.43	1.00		
Early	.50	.48	.35	.94	1.00	
Late	.47	.36	.43	.88	.66	1.00

[a] Between sample variation in correlations significant ($p < .05$) by Hays' (1963, p. 532) V.

GENERAL DISCUSSION

Eight URI symptoms adequately measured this illness in basic training. These symptoms -- fever, sore throat, dry cough, productive cough, stuffed-up nose, sneezing, hoarseness, and sinus pain -- clearly involved the upper respiratory tract. With the exception of hoarseness, these symptoms are among the most commonly reported URI symptoms in studies of recruit and civilian populations (Forsyth, et al., 1963; Gwaltney, et al., 1967; Verbrugge, 1986). The eight symptoms, therefore, are reasonably representative of common colds as they occur in a variety of settings.

Constitutional symptoms were conspicuously absent from the final URI measure. Malaise and myalgia, two symptoms commonly used in URI assessments, occurred frequently but were more likely to be associated with musculoskeletal problems than with URI in basic training. These findings should be of concern whenever young, generally healthy populations are studied, because musculoskeletal problems are the major non-URI source of illness in this age group in the general U. S. population (Verbrugge, 1986). Other constitutional symptoms commonly found in URI composites (e.g., anorexia, nausea) occurred infrequently and were not specifically associated with URI when they did occur. The low frequency of these symptoms was consistent with other surveys of military personnel (Forsyth, et al., 1963). These infrequent constitutional symptoms may indicate other, relatively infrequent, types of infection but are of limited importance for measuring URI. Here again, the results may generalize beyond military populations, because low frequencies of gastrointestinal symptoms have been reported in surveys of civilian populations (Gwaltney, et al., 1967).

Psychometric evaluation indicated that the eight URI symptoms defined an illness dimension, rather than a discrete class variable. Location on this dimension is measured best by average symptom severity adjusted for the presence of allergy and musculoskeletal problems. This assessment approach is consistent with recommended clinical practice and provides an index that combines satisfactory reliability with minimal overlap with general symptom reporting. Note also that the recommended scoring procedure reduces to a simple symptom severity score comparable to other URI measures in the literature whenever allergy and musculoskeletal problems are not reported.

The overlap between the proposed URI measure and general symptom reporting was large enough to pose interpretational problems. Noting that

the symptoms of vomiting and diarrhea might represent non-URI infections, some of the overlap may represent general susceptibility to infections. As a check on this possibility, the correlations between URI and the present indicator of general symptom reporting were compared to those obtained with a composite of constipation, dizziness, sunburn, and skin irritation. The correlations were virtually identical. However, given the fact that viral infections can produce such a wide range of symptoms and considering the possibility of secondary effects, it is impossible to absolutely rule out an infection interpretation of any symptom composite. For example, constipation and dizziness may be occasional secondary effects of dehydration arising from a fever. The important point is that, despite this ambiguity, the substantial empirical overlap between indicators of general symptom reporting and URI should be a concern when interpreting the results of studies examining correlates of URI. One approach to dealing with this problem is to include assessments of non-URI health problems in studies when URI is the focal illness. Associations between URI and other variables can be compared to the associations between those other variables and other types of illness. Such comparisons should help decide whether an observed association should be interpreted as evidence that a predictor is related to disease or to general symptom reporting tendencies which may be purely psychological in character. Prior studies of military recruits have shown this to be a useful strategy (Vickers, 1986; Vickers, Hervig & Edwards, 1986).

The URI measure developed in this study provides a simple method for evaluating this illness in military populations. Other research evidence that symptom reports correlate with alternative measures of URI, including viral shedding and clinical evaluations by medical personnel, provides a basis for asserting that the symptom report measure described here can be accepted as a valid indicator of infection. Given the invasive, costly nature of the alternatives, symptom reports provide an efficient method of testing hypotheses to identify issues for later detailed study and confirmation by alternative methods.

REFERENCES

Anderson, L.J., Patriarca, P.A., Hierholzer, J.C. & Noble, G.R. (1983). Viral respiratory illnesses. Medical Clinics of North America, 67, 1009-1030.

Arlander, T.R., Pierce, W.E., Edwards, E.A., Peckinpaugh, R.O. & Miller, L.F. (1965). IV. An epidemiologic study of respiratory illness patterns in Navy and Marine Corps recruits. American Journal of Public Health, 55, 67-80.

Beare, A.S. & Reed, S.E. (1977). The study of antiviral compounds in volunteers. In J. Oxford (ed.), Chemoprophylaxis and virus infections of the respiratory tract. Volume 2. Cleveland: CRC Press, pp. 27-55.

Brown, M.B. (1983). Two-way and multiway frequency tables -- measures of association and the log-linear model (complete and incomplete tables). In Dixon, W.J. (ed.), BMDP Statistical Software. Berkeley, CA: University of California Press, pp. 143-206.

Cohen, J. (1978). Partialled products are interactions; partialled powers are curve components. Psychological Bulletin, 85, 858-866.

Dingle, J.H., Badger, G.F. & Jordan, W.S., Jr. (1964). Illness in the Home: A Study of 25,000 Illnesses in a Group of Cleveland Families. Cleveland: Press of Western Reserve University.

Dowling, H.F. & Lefkowitz, L.B. (1963). Clinical syndromes in adults caused by respiratory viruses. American Review of Respiratory Disease, 88, 61-72.

Edwards, E.A. & Rosenbaum, M.J. (1971). The Surveillance Program -- 1964-1970. Great Lakes, IL: Naval Medical Research Unit No. 4, Report 71.8.

Forsyth, B.R., Bloom, H.H., Johnson, K.M. & Chanock, R.M. (1963). Patterns of illness in rhinovirus infections of military personnel. New England Journal of Medicine, 269, 602-606.

Gorsuch, R.L. (1984). Factor Analysis. Hillsdale, NJ: Erlbaum.

Gwaltney, J.M., Jr., Hendley, J.O., Simon, G. & Jordan, W.S., Jr. (1967). Rhinovirus infections in an industrial population. II. Characteristics of illness and antibody response. Journal of the American Medical Association, 202, 158-164.

Harlan, W.R., Murt, H.A., Thomas, J.W., Lepkowski, J.M., Guire, K.E., Berki, S.E. & Landis, J.R. (1986). Incidence, utilization, and costs associated with acute respiratory conditions, United States, 1980. National Medical Care Utilization and Expenditure Survey. Series. C, Analytical Report No. 4. DHHS Pub. No. 86-20404. National Center for Health Statistics, Public Health Service. Washington. U. S. Government Printing Office, Sept., 1986.

Hays, W.L. (1963). Statistics for Psychologists. NY: Holt, Rinehart, Winston.

Heise, D.R. (1969). Separating reliability and stability in test-retest correlation. American Sociological Review, 334, 93-101.

Jackson, G.G., Dowling, M.D., Spiesman, M.O. & Boand, A.V. (1958). Transmission of the common cold to volunteers under controlled conditions. I. The common cold as a clinical entity. Archives of Internal Medicine, 101, 267-278.

Jackson, G.G. & Muldoon, R.L. (1973). Viruses Causing Common Respiratory Infections in Man. Chicago: University of Chicago Press.

Lowenstein, S.R. & Parrino, T.A. (1987). Management of the common cold. Advances in Internal Medicine, 32, 207-233.

Lytle, R.I. & McNamara, M.J. (1967). Alpha-1 globulins in those susceptible to viral acute respiratory diseases. Great Lakes, IL: Naval Medical Research Unit No. 4, December, 1967.

Lytle, R.I., Rytel, M.W. & Edwards, E.A. (1966). Correlation of immunological and biological factors in the host with susceptibility to respiratory infections. Journal of Infectious Diseases, 116, 67-74.

Meehl, P.E. (1973). Psychodiagnosis: Selected Papers. Minneapolis, MN: University of Minnesota Press, pp. 200-244.

Monto, A.S., Napier, J.A. & Metzner, H.L. (1971). The Tecumseh study of respiratory illness. I. Plan of study and observations on syndromes of acute respiratory disease. American Journal of Epidemiology, 94, 269-279.

National Center for Health Statistics. (1983). Current Estimates from the National Health Interview Survey, United States, 1983. Vital and Health Statistics. Series 10, No. 154. DHHS Pub. No. (PHS) 86-1582. National Center for Health Statistics, Public Health Service. Washington. U. S. Government Printing Office.

National Center for Health Statistics. (1986). Current Estimates from the National Health Survey. United States, 1983. Washington, D.C.: U. S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, Series 10, No. 154.

Nunnally, J.C. (1978). Psychometric Theory. NY: McGraw-Hill.

Roden, A.T. (1958). Clinical assessment of the common cold. Proceedings of the Royal Society for Medicine, 51, 271-273.

Roden, A.T. (1963). Variations in the clinical pattern of experimentally induced colds. Journal of Hygiene, 61, 231-246.

Rosenthal, R. (1978). Combining the results of independent studies. Psychological Bulletin, 85, 185-193.

Thomas, G.B., Jr. (1972). Calculus and Analytical Geometry. Reading, MA: Addison-Wesley, pp. 1-158.

Totman, R., Kiff, J., Reed, S.E. & Craig, J.W. (1980). Predicting experimental colds in volunteers from different measures of recent life stress. Journal of Psychosomatic Research, 24, 155-163.

Totman, R., Reed, S.E. & Craig, J.W. (1977). Cognitive dissonance, stress and virus-induced common colds. Journal Of Psychosomatic Research, 21, 55-63.

Verbrugge, L. (1986). From sneezes to adieux: stages of health for American men and women. Social Science and Medicine, 22, 1195-1212.

Vickers, R.R., Jr. (1986). Psychosocial risk factors for upper respiratory infections: Demographic predictors of outpatient treatment. San Diego: Naval Health Research Center, Technical Report 86-31.

Vickers, R.R., Jr., Hervig, L.K. & Edwards, E.A. (1986). Psychosocial risk factors for upper respiratory infection: An exploratory study. San Diego: Naval Health Research Center, Technical Report 86-27, 1986.

Vickers, R.R., Jr., Hervig, L.K., Edwards, E.A. & Rahe, R.H. (in preparation). Emotion, immune system serum proteins, and health status in a stressful situation.

Appendix A

APPLICATION OF MEEHL'S TEST FOR DISCRETE CATEGORIES TO URI ASSESSMENT

General Background

Common experience suggests that upper respiratory infections produce illnesses of varying severity, but symptom reports may be useful primarily for distinguishing people who have an infection from those who do not. Meehl's (1973) MAXCOV test for discrete categories was applied to determine whether URI measures should be scored to differentiate between two discrete categories of recruits, healthy and ill, or to differentiate between locations on a continuum reflecting severity of illness. The logic of the MAXCOV test and a bootstrap procedure to test its appropriateness are described below. The bootstrap procedure then is applied to the URI data.

Conceptual Basis for the Approach

The MAXCOV procedure can be applied if there is a set of URI symptoms that are more frequent among recruits with upper respiratory infections than among those without such infections. If the discrete category model holds, the covariance between any two symptoms in this set can be expressed as:

$$cov(xy) = [p*cov_h(xy)] + [q*cov_i(xy)] + [pq*(x_i-x_h)*(y_i-y_h)] \quad \text{(Equation 1)}$$

where "cov" indicates covariance, "x" and "y" indicate two distinct symptoms, "h" and "i" refer to healthy and ill categories, and "p" and "q" refer to proportions of healthy and ill individuals with q = (1-p). The subscripted "x" and "y" values in the equation indicate the group means for the variables under consideration. Note that the last term in the equation must be positive, given the assumption that the symptoms are more common in the ill group.

If the hypothesized healthy and ill groups have equal within group covariances, Equation 1 reduces to:

$$cov(xy) = (p+q)*k_1 + (p*q*k_2)$$

$$= k_1 + (p*q*k_2) \quad \text{(Equation 2)}$$

where k_1 is the within-group covariance and k_2 is the covariance attributable to differences in symptom reporting rates for the two categories of recruits. The (p+q) term equals 1.00, because all recruits must be either infected or healthy.

Bootstrap Test for Discrete Categories

Equation 2 leads to a bootstrap test of the hypothesis of discrete categories based on the observation that cov(xy) will vary systematically as a function of the product "p*q" if the hypothesis is true. In particular, the covariance reaches a minimum value of k_1 when all recruits are healthy or when all are infected, because "p*q" equals zero. The maximum value, $k_1 + .25k_2$, occurs when $p = q = .50$.

Given symptoms that are more common among ill recruits, groups can be defined on the basis of total score on a composite of those symptoms. The proportion of ill recruits in each group should increase when the groups are ordered by total score. The bootstrap test of the model, therefore, utilizes the number of symptoms reported for a subset of the total symptoms to define groups, estimates the within group covariance for symptoms in each group, then checks for the presence of a peaked covariance curve. The bootstrap procedure is as follows:

(a) Set aside two symptoms to serve as x and y.

(b) Sum the remaining symptoms to produce a URI symptom subset composite score. For these analyses, each symptom was recoded to indicate a simple presence-absence distinction to reduce the number of possible total scores sufficiently to ensure sample sizes adequate to compute reasonably precise computations of covariance within each group. As a result, symptom subset scores could range from 0 to 6.

(c) Classify subjects into groups defined by their scores on the subset.

(d) Compute the sample covariance between x and y for each of the seven groups defined by the symptom subset.

(e) Repeat (a) through (d) with a new pair of symptoms until all possible symptom pairs have been considered. Given 8 symptoms, there were 28 such pairs.

(f) Sum the covariance estimates for each group level across all 28 symptom pairs. For example, add the estimates for the groups defined by the score "0" in each analysis.

(g) Determine whether the summed covariances produce the peaked covariance curve predicted by Meehls' (1973) model.

Application of MAXCOV to URI

The MAXCOV procedure was applied to the data from the first through fourth and the sixth sessions for each of the four samples, thereby producing 20 separate analyses. Data from the sixth session were included to have one day with a relatively low expected rate of URI as a basis for evaluating the possibility that results obtained with data from earlier sessions were biased by high base rates of illness.

Results were tabulated separately for each day in each of the four studies (Table A-1). This approach was adopted, because the proportion of healthy and ill recruits was likely to vary across sessions and across samples. If so, different session-study combinations would have graphs which peaked at different points. Averaging across session-sample combinations therefore could produce a misleading flat graph.

Results

The assumption of equality of the within group covariances for healthy and ill individuals was tested as the first step in the analysis. The extreme groups were compared as the best indicators of "pure" healthy and ill groups. The covariance was higher in the healthy group in 14 of 20 analyses (chi-square = 5.00, $p < .051$). Hartley's F_{max} test with an assumed 27 degrees of freedom for each covariance estimate indicated the covariance in the healthy group was significantly ($p < .05$) larger in two of the 20 analyses. This many significant findings could be expected to occur 26% of the time in a series of 20 analyses. On the whole, these analyses offered no reason to reject the assumption of equal within group covariances.

One-way analyses of variance tested for the peaked covariance curve predicted by the MAXCOV model. The classification variable for these

analyses was the total score on the symptom subsets. The dependent variable was the within-group covariance estimate derived from the bootstrap procedure for that score level. As noted previously, there were 28 such estimates for each score level, one for each distinct subset of symptoms. Each of the covariance estimates was treated as a separate within-cell case. The one-way ANOVA tested for group differences, including a polynomial trend analysis to test for the presence of a peaked curve.

Table A-1

Summary of MAXCOV Analyses

	Group Classification Score:							ANOVA F-tests:			
Session	0	1	2	3	4	5	6	Over[a]	Lin[b]	Quad[c]	High[d]
Study A											
1	1186	1855	1283	2782	3777	697	1348	6.22	.08	11.92	6.32
2	1197	2625	2164	859	3672	620	918	8.02	2.49	8.01	9.41
3	1464	2435	171	4534	975	1896	827	12.39	1.03	7.39	16.48
4	822	2841	1314	542	2799	1919	1519	4.18	.56	.48	6.01
6	991	1221	2637	3363	526	1103	2578	4.60	.84	1.25	6.38
Study B											
1	861	1887	1888	3314	2591	1403	533	4.63	.28	23.07	1.11
2	1450	-110	6167	473	2219	1268	479	21.87	2.95	17.71	27.64
3	2381	1115	2051	1593	3683	127	1151	6.59	3.00	2.15	8.59
4	1080	2304	1342	1928	1946	1878	674	2.14	.49	5.94	1.60
6	809	1851	1196	4946	-1013	4760	452	23.77	1.10	11.24	32.57
Study C											
1	899	1321	3324	3120	1007	2543	440	6.79	.28	21.01	4.86
2	1342	2444	-282	5132	773	1845	672	16.11	.87	8.95	21.71
3	856	2542	1670	1571	2453	1394	1091	2.19	.13	5.05	1.98
4	1156	1836	814	5088	459	2307	936	9.41	.00	8.43	12.01
6	415	2985	1233	854	4429	1981	848	10.29	1.10	11.77	12.22
Study D											
1	955	1219	3863	1477	950	2969	932	7.74	.05	7.98	9.60
2	1774	1843	405	3698	1350	1049	1438	6.57	.61	1.19	9.40
3	1756	2895	-658	4247	813	781	2043	12.48	.63	.14	18.53
4	1050	2058	1067	1512	2507	2913	277	5.35	.15	7.71	6.06
6	1161	1384	2041	961	2718	2956	1242	2.69	2.49	1.87	2.95

[a] Overall F-test, 6,189 df, $p < .05$ if $F > 2.15$, $p < .01$ if $F > 2.90$

[b] Linear trend with 1,189 df, $p < .05$ if $F > 3.90$, $p < .01$ if $F > 6.78$

[c] Quadratic trend with 1,189 df, $p < .05$ if $F > 3.90$, $p < .01$ if $F > 6.78$

[d] Higher order trends with 4,189 df; $p < .05$ if $F > 2.43$, $p < .01$ if $F > 3.42$

Significant between group differences in covariance were obtained in all 20 analyses (4 groups with 5 days data for each group). The quadratic trend was significant ($p < .05$) in 14 of 20 cases. This finding was important, because the well-defined peak in the covariance curve predicted by the MAXCOV procedure would manifest itself empirically as a quadratic trend. However, higher order trend components were substantial and significant in 17 of 20 analyses. This finding implied that a simple single-peaked curve would not describe the functional relationship between total score and covariance (Thomas, 1972). Direct examination of the covariances suggested that in many instances there were multiple peaks of comparable magnitude. Thus, the overall results of the analysis failed to confirm the presence of a single-peaked covariance curve, despite the significant quadratic trend.

Conclusion

The data did not fit the MAXCOV model. Insofar as this model provides an adequate test for discrete groups, the data contradicted the hypothesis that such groups existed in the population studied. Therefore, it was concluded that the symptoms defined a continuum of illness severity.

Appendix B

MEASUREMENT OF GENERAL SYMPTOM REPORTING

The measure of general symptom reporting employed in these analyses consisted of responses to symptom checklist items concerning skin irritation, vomiting, diarrhea, and trouble hearing. These four symptoms were selected from those which did not produce consistent group orderings when healthy recruits were compared to those complaining of musculoskeletal problems and colds (see Table 2, p. 7). These symptoms were selected from among those which did not involve the respiratory system. An attempt was made to include symptoms occurring in different body sites, if not precisely in different body systems.

After selection, there was some concern that vomiting and diarrhea might represent part of a gastrointestinal infection syndrome, but the correlations between these symptoms was quite low. In fact, each of these symptoms typically correlated more strongly with either skin irritation or trouble hearing than with one another.[1]

The four symptoms selected clearly defined a single factor as principle components analysis showed a strong first factor in each of 16 analyses (median eigenvalue = 1.60, range = 1.26 - 1.90) with no other eigenvalue as large as 1.00 in any analysis. The average interitem correlation ranged from .08 to .30 (median = .21) with only 12 of 96 interitem correlations larger than .30. In addition, the symptoms were relatively rare in this population as mean scores for the composite ranged from 1.13 to 1.34.

[1]As a check on the most important possible effect of an inappropriate selection of symptoms, a second composite, consisting of constipation, dizziness, sunburn, and skin irritation was constructed. These symptoms were selected because they were the ones in the checklist that seemed least likely to derive directly from viral infections. The correlation between the URI measures and this alternative composite then were computed and compared to those for the composite described above. The associations were virtually identical, so the specific composite employed probably did not substantially affect the estimates of URI-General Symptom Reporting associations. It should be noted, however, that it is very difficult to be certain that any general symptom composite does not contain indicators of viral infection. This is partly because viral infections can involve nearly every body system and partly because some symptoms may arise as secondary effects of primary viral symptoms. For example, dizziness and constipation could be secondary effects of fever.

The analysis findings suggested that high scores on the composite could legitimately be regarded as evidence of a general tendency to endorse symptoms. One reason was the interitem correlations generally were moderate for all pairs of items, as would be expected from the effects of a general response tendency. A second reason was that the relative rarity of these symptoms in the population made it more likely that high scores would occur only in people with a marked tendency to endorse a wide range of symptoms. Finally, none of the symptoms had a frequent occurrence in conjunction with URI, so there was virtually no likelihood that the composite would confound two measures of the same construct.

Appendix C

COMPARISON OF CORRELATIONS FOR URI MEASURES FROM INITIAL AND LATER SESSIONS

The temporal pattern of URI scores was investigated because evidence from laboratory studies of viral infections suggested that the typical cold produces symptoms for one or two weeks, although some lingering colds produced symptoms for up to four weeks (Roden, 1958). Given this evidence, valid URI measures obtained at one point in training would be expected to correlate moderately with later measures of URI. The strength of the association should decrease over time but may last up to four weeks.

URI measures obtained at one session generally were moderately correlated with those obtained one session later and smaller correlations were obtained for measures with longer time intervals between assessments (Table C-1). When sessions were four weeks apart, the correlation between the earlier and later URI approximated the correlation between the earlier URI and general symptom reporting at the later point in time. This latter correlation was included as a reference point to determine the effects of general symptom reporting tendencies. Therefore, the general trends in the findings were consistent with expectations derived from laboratory studies of URI.

Table C-1

Cross-time Correlations for URI Severity

	Session 1		Sessions 2-7			
	URI	General	URI		General	
Lag	r	r	Mean	S.D.	Mean	S.D.
URI Severity						
1	.38	.35	.59	.02	.37	.04
2	.29	.30	.45	.04	.34	.02
3	.25	.28	.38	.05	.32	.04
4	.25	.28	.32	.01	.30	.02

NOTE: Each correlation considered in these analyses related measures taken at one session to measures taken in later sessions. "Lag" was the difference between the session number for the later measure and the session number for the earlier measure. For example, the correlation between measures from sessions 1 and 2 would have a lag of 1, the correlation between measures from 2 and 5 would have a lag of 3, and so on. Thus, longer lags imply more time between measures with each lag increment representing approximately a 1 week interval.

The stability coefficients for the URI score derived from data collected in the first session deviated noticeably from the general pattern. These coefficients were substantially lower than the average lagged correlations for the later sessions and were closely comparable to the corresponding General Symptom correlations for each lag interval.

The data from the initial session were excluded from subsequent analyses, because their pattern of association to later URI measures deviated from the pattern for the later sessions. The major consideration in this decision was that the results obtained with these data would not be representative of the typical performance of the URI assessment. This decision did not represent a judgment that the initial session produced invalid data, only that it produced unrepresentative data. The basis for the difference between this session and later sessions remains to be determined. Further study of this point could be important if the differences imply that URI reports are sensitive to situational factors such as initial exposure to a symptom report instrument or acute psychological reactions to a stressful setting.

Appendix D

PSYCHOMETRIC COMPARISONS OF ALTERNATIVE URI MEASURES

The psychometric characteristics of three alternative symptom measures for URI were evaluated. Ideally, an indicator of URI should have high measurement precision, moderate stability over time (consistent with the transient nature of URI), and minimal contamination by generalized symptom reporting tendencies which might arise from acute stress reactions or general hypochondriacal tendencies. This appendix describes analyses comparing these attributes for URI Severity, URI Symptoms, and Adjusted URI Severity.

Analysis Procedures. Heise's (1969) path analytic approach to estimating reliability and stability of scales was applied in the first phase of the analyses. This procedure was chosen because it could be applied to the measure of General Symptom reporting. Typical procedures based on internal consistency seemed inappropriate for this measure as strong correlations between the items in this composite were not necessarily expected.

The first phase of the analysis produced reliability and stability estimates. Given data from three sessions and designating correlations between scores for different sessions as r_{ij}, the formulae for computing reliability (r_{xx}) and stability from session i to session j (s_{ij}) are:

$$r_{xx} = (r_{12} * r_{23})/r_{13}$$
$$s_{12} = r_{13}/r_{23}$$
$$s_{23} = r_{13}/r_{12}$$

The second phase of the analyses employed corrected URI-General Symptom correlations for attenuation due to measurement error. The corrected values were the observed correlations divided by the square root of the product of the estimated reliabilities (Nunnally, 1978, pp. 219-220).

The first two phases of the analyses were conducted separately for each sample. The third phase estimated pooled coefficients and between-sample variation in the findings. Each coefficient computed in the first two phases represented a correlation, because reliability coefficients estimate the correlation between measured and true scores (Nunnally, 1978, pp.

193-203). Hays (1963, pp. 529-533) has provided a test for the hypothesis that correlations obtained in independent samples are equal. The formula for his V statistic is:

$$V = (n_j-3)\ (z_j-U)^2/\ \ (n_j-3),$$

where z_j is the Fisher r-to-z transformation of the correlation for sample j, and n_j is the sample size in the jth sample (Hays, 1963, p. 533). U is the weighted average of z_j for the samples involved, computed by multiplying each transformed correlation by the associated n_j, then dividing by the sum of the n_j. V is distributed as a chi-square with (k-1) degrees of freedom, where k is the number of correlations compared.

The weighted average coefficients and associated chi-square information from the third phase are shown in Table C-1. Chi-squares have been expressed as the ratio of each chi-square to its degrees of freedom. In related types of analyses, a ratio between 2.00 and 3.00 is generally accepted as satisfactory fit between the model and the data even when this ratio implies an overall chi-square which is, technically, statistically significant.

Table D-1

Comparison of Alternative URI Measures

	Symptom Severity		Symptom Count		Adjusted Severity	
	r (rms)	x^2/df	r (rms)	x^2/df	r (rms)	x^2/df
Early Period						
URI Reliability	.843 (.010)	0.58	.759 (.045)	6.17	.798 (.046)	8.74
URI Stability 1-2	.671 (.036)	2.14	.709 (.041)	3.74	.617 (.057)	4.63
URI Stability 2-3	.728 (.049)	4.93	.810 (.073)	15.09	.662 (.076)	11.60
URI-General Symp.	.582 (.050)	1.95	.566 (.051)	1.55	.516 (.071)	2.99
Gen. Reliability	.743 (.030)	2.16	.680 (.024)	3.07	.743 (.029)	2.16
Gen. Stability 1-2	.794 (.096)	40.80	.852 (.081)	65.69	.794 (.096)	40.80
Gen. Stability 2-3	.871 (.029)	7.22	.896 (.117)	142.15	.871 (.029)	7.22
Late Period						
URI Reliability	.744 (.088)	11.27	.716 (.061)	3.89	.669 (.085)	6.69
URI Stability 6-7	.833 (.105)	54.17	.828 (.074)	15.40	.755 (.094)	17.52
URI Stability 7-8	.826 (.099)	20.50	.873 (.066)	18.21	.779 (.131)	23.74
URI-General	.746 (.167)	39.51	.755 (.143)	15.75	.610 (.198)	20.80
Gen. Reliability	.612 (.055)	2.16	.607 (.069)	3.34	.612 (.055)	2.16
Gen. Stability 6-7	.939 (.106)	169.98	.942 (.177)	228.43	.939 (.106)	169.98
Gen. Stability 7-8	.765 (.133)	26.61	.826 (.120)	21.44	.765 (.133)	26.61

NOTE: See text for description of computational procedures. Table entries are average coefficient (root mean square for deviations from average) and chi-square/degrees of freedom ratio.

Results. The analyses were applied to data from three sessions early in training and three late in training, excluding session 1 for reasons given in Appendix B. The $\underline{V}$ statistics for the reliability and stability coefficients for the URI measures frequently were well outside the recommended range for the chi-square goodness-of-fit criterion (Table D-1). Note, however, that the goodness-of-fit for the URI measures generally was better than the corresponding value for the General Symptom reports.

Examination of the differences between the average value and the values for individual studies showed that Study B was much more deviant than any of the remaining samples. The overall goodness-of-fit was substantially improved by deleting Sample B, but a number of the chi-square/df ratios still were large enough to be a reason for concern for each URI measure (Table D-2). Thus, the sample-to-sample variability was not attributable to the apparent outlier sample.

Given substantial variability across samples, a sample-by-sample comparison of the alternative URI measures was made. In this comparison, the magnitudes of the coefficients were compared within each sample, and the results of these comparisons were aggregated across samples (Table D-3). The results can be summarized as follows:

Reliability: Severity > Symptoms = Adjusted Severity
Stability: Symptoms > Severity > Adjusted Severity
U-G Correlation: Severity = Symptoms > Adjusted Severity

The consistency of these pairwise comparisons across the four samples was more important than the variation in the magnitude of the associations across samples. This conclusion was reasonable, because the relative size of coefficients for different measures was the basis for comparing the measures.

Table D-2

Comparison of Alternative URI Measures with Sample B Deleted

	Symptom Severity		Symptom Count		Adjusted Severity	
	r (rms)	x^2/df	r (rms)	x^2/df	r (rms)	x^2/df
Early Period						
URI Reliability	.837 (.011)	0.62	.765 (.045)	7.90	.787 (.046)	11.83
URI Stability 1-2	.663 (.036)	3.08	.703 (.041)	5.55	.612 (.057)	6.94
URI Stability 2-3	.729 (.049)	7.18	.794 (.072)	18.96	.681 (.081)	13.32
URI-General	.571 (.051)	1.13	.557 (.051)	1.52	.514 (.071)	2.83
Gen. Reliability	.742 (.035)	3.16	.683 (.024)	1.20	.742 (.029)	3.16
Gen. Stability 1-2	.735 (.104)	11.91	.789 (.089)	5.58	.735 (.104)	11.91
Gen. Stability 2-3	.860 (.030)	7.43	.828 (.110)	34.73	.860 (.030)	7.43
Late Period						
URI Reliability	.780 (.100)	8.04	.748 (.072)	0.30	.704 (.096)	4.97
URI Stability 6-7	.745 (.111)	3.33	.802 (.075)	13.60	.695 (.104)	3.40
URI Stability 7-8	.783 (.099)	24.59	.847 (.066)	12.55	.727 (.129)	14.62
URI-General	.621 (.173)	6.08	.711 (.142)	18.84	.493 (.213)	8.55
Gen. Reliability	.640 (.064)	0.74	.642 (.080)	1.17	.640 (.064)	0.74
Gen. Stability 6-7	.869 (.094)	28.16	.870 (.155)	116.09	.869 (.094)	28.16
Gen. Stability 7-8	.799 (.145)	30.81	.820 (.118)	31.27	.799 (.145)	30.81

NOTE: See text for description of the computational procedures. Table entries are average coefficient (root mean square for deviations from average) and chi-square/degrees of freedom ratio.

Table D-3

Summary of Sample-by-Sample Comparison of Alternative URI Measures

		Number of Comparisons for Which: Severity> Symptoms	Severity> Adjusted Severity	Symptoms> Adjusted Severity
Early	R_{xx}	4	3	1
	S_{12}	0	4	4
	S_{23}	0	3	4
	U-M	4	12	9
Late	R_{xx}	3	4	4
	S_{56}	1	4	4
	S_{67}	1	4	4
	U-M	5	12	9
Chi-Square	R_{xx}	4.50	4.50	0.50
	Stab.	9.00	12.25	16.00
	U-M	0.33	24.00	6.00

NOTE: R_{xx} = Reliability coefficient; S_{xy} = Stability coefficient for the two sessions indicated by the numbers. Each of the three daily estimates for the URI-Musculoskeletal correlation, designated U-M, was considered separately. Chi-squares were computed for the combined early-late periods collapsing across all comparisons for the indicated coefficient.

Tables D-1 through D-3 make it clear that the average values for the coefficients represent those patterns accurately, whether the sample most deviant from the average was included or not. Therefore, the average values provided reasonable descriptions of general trends in the data and a suitable illustration of the bases employed to choose between the alternative methods.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION (U) UNCLASSIFIED	1b RESTRICTIVE MARKINGS None
2a. SECURITY CLASSIFICATION AUTHORITY N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NHRC Report No. 88- 23	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Health Research Center 6b. OFFICE SYMBOL (If applicable) 70	7a. NAME OF MONITORING ORGANIZATION Commander, Naval Medical Command
6c. ADDRESS (City, State, and ZIP Code) P.O. Box 85122 San Diego, CA 92138-9174	7b. ADDRESS (City, State, and ZIP Code) Department of the Navy Washington, DC 20372
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Naval Medical Research & Development Command 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) Naval Medical Command National Capital Region Bethesda, MD 20814-5044	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO.: MR04101 / PROJECT NO.: .07 / TASK NO.: 6004 / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification) (U) PSYCHOSOCIAL RISK FACTORS FOR UPPER RESPIRATORY INFECTIONS: Assessment of Upper Respiratory Illness during Basic Training

12. PERSONAL AUTHOR(S)
Vickers, Ross R., Jr. & Hervig, Linda K.

13a. TYPE OF REPORT	13b TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM _____ TO _____	1988 May 23	

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	health, symptoms, measurement, recruits, upper respiratory infection

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Symptom reports have been useful for epidemiological surveillance of upper respiratory illness (URI) in civilian populations, but available measures may be ill-suited to military populations performing hard physical work under psychologically stressful conditions. Thirty-nine potential URI symptoms were evaluated in four samples of Navy recruits to develop a URI symptom report measure suitable for such circumstances. Sore throat, productive cough, stuffed-up nose, dry cough, sinus pain, sneezing, fever, and hoarseness were retained based on their relationships to reported health status and correlations between symptoms. Scores on the composite of these 8 symptoms correlated with reports of concurrent musculoskeletal problems and allergy, so adjustments for concurrent illnesses were incorporated. The resulting measure is suitable for epidemiological surveillance of URIs in military personnel, provided proper allowance is made for the influence of general symptom reporting tendencies.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Vickers, Ross R., Jr.	22b TELEPHONE (Include Area Code) 619-553-8454 22c. OFFICE SYMBOL 70